FILE TAGGING AND AUTOMATIC CONVERSION OF DATA OR FILES

ABSTRACT OF THE DISCLOSURE

The present invention provides facilities for tagging files or data with attribute information in the form of a file tag (TAGINFO) which contains an identifier for text information (TXTFLAG) and an attribute (CCSID) for identifying encoding schemes. TXTFLAG is an auto conversion flag that inhibits automatic conversion between encoding schemes when switched off, while CCSID is an encoding scheme identifier. Furthermore, a runtime attribute (process CCSID) is assigned to a process specifying the runtime encoding scheme. A conversion is done automatically by an auto conversion function if both CCSIDs allow a conversion. Files having no file tag are tagged with a virtual file tag (default tag) by means of an automatic tagging (AUTOTAG) function using heuristic rules for determining whether the data or file contains text or binary information. Old applications must work with untagged files as before. Existing applications should be able to benefit from auto conversion and thereby to be enabled to process new, tagged files without code changes. This invention allows to physically store data in the process codepage of the application thereby avoiding any conversions in the frequently used path while the file tagging and auto conversion does not inhibit other programs running in a different codepage to access the data.